

חיזוי סיכון להפרעת דחק פוסט טראומטית  
באמצעות גורמים אקולוגיים בלבד

יישום של למידה עמוקה

# רקע

## הפרעת דחק פוסט טראומטית [1]

---

- פורצת כתוצאה מחשיפה לאירוע טראומטי
- באה לידי ביטוי בסימפטומים ובמצוקה פסיכולוגית חזקה: *Reexperiencing*, *Avoidance, & Hyperarousal*
- חשיבות של זיהוי אינדיבידואלים בסיכון גבוה להפרעה טרם חשיפה לאירוע טראומטי



# אטיולוגיה [4,5,7,9]

## Diathesis-stress model of PTSD

- Severity of trauma
- Biological diathesis
- Ecological diathesis

- מחקרים זיהו גורמי סיכון רבים להפרעת דחק פוסט טראומטית
- אך אין לממצאים משמעות פרקטית קלינית עבור הפרט
- נובע מהשימוש במודלים סטטיסטיים שנועדו לבחון הבדלים בין קבוצות ולא לתת חיזוי עבור הפרט עם נתונים חדשים שהמודל לא ראה.

# Machine Learning [4,5]

---

- **אתגר חישובי** לתת חיזוי עבור הפרט. צריך גישה אחרת
- האינטראקציה בין גורמי הסיכון השונים להפרעת דחק פוסט טראומטית, מורכבת, יש תלות מותנית בין גורמי הסיכון. על כן יש יחס מורכב בין **הקלט** (גורמי סיכון) ו**הפלט** (סיכון להפרעת דחק פוסט טראומטית)
- הגישה של למידת מכונה מתאימה

# מחקרים קודמים [4,5,9]

- מחקרים קודמים השתמשו ב-*SVM* לחיזוי הפרעת דחק פוסט טראומטית בהתבסס על **ביו-מרקרים** שונים שנלקחו מסריקות MRI, fMRI ו PET
- מחקרים **מועטים מאוד** השתמשו בנתונים שאינם ביו-מרקרים לחיזוי סיכון להפרעת דחק פוסט טראומטית. **אין מחקר** שהשתמש ברשימה מקיפה של **גורמים אקולוגיים** לחיזוי סיכון להפרעת דחק פוסט טראומטית.
- מדוע דווקא גורמים אקולוגיים? **מסלול אטיולוגי שנזנח לחלוטין במחקר למידת מכונה** – קל לאסוף נתונים (מחיר)
- גלזר-לוי, שלו, וקרסטופט (2014) יישמו *SVM* לחיזוי נבדקים שפיתחו הפרעת דחק פוסט טראומטית לאחר חוויה של אירוע טראומטי. הנתונים שנאספו במחקרם, היו נתונים דמוגרפיים, לחץ דם, דופק, רמת כאב, סוג האירוע, סוג הפגיעה
- קרסטופט, סטטניקוב, אנדרסן, מדסן וגלזר-לוי (2015) יישמו *SVM* על נתונים **דמוגרפיים** שונים כדוגמת מין, מצב משפחתי, צריכת אלכוהול, רמת השכלה וכו' על מנת לחזות חיילים דנים שפיתחו הפרעת דחק פוסט טראומטית



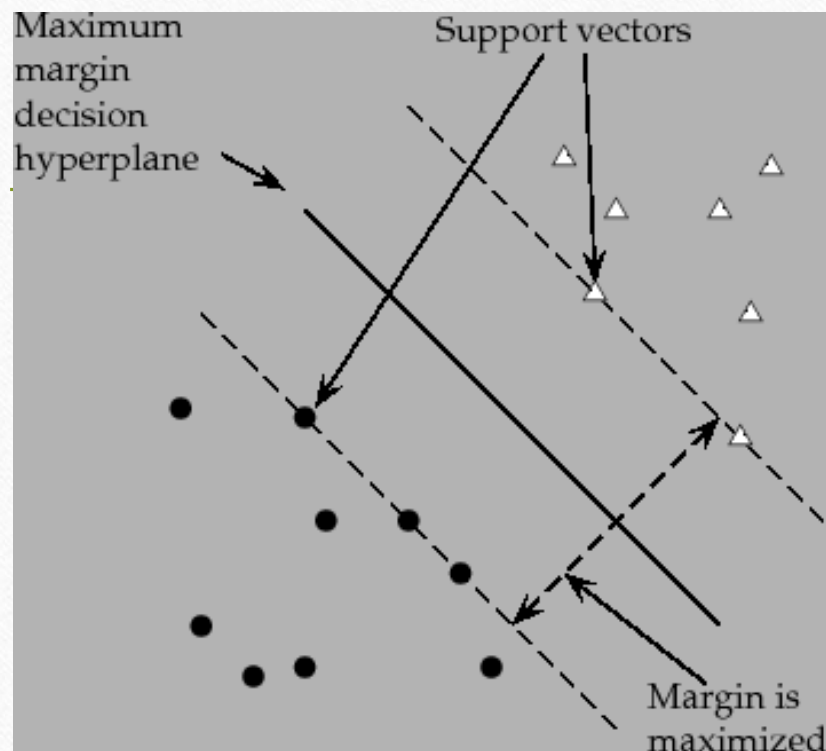
# מטרות הפרויקט סקירה כללית

- בניית מודל לחיזוי סיכון להפרעת דחק פוסט טראומטית עבור הפרט באמצעות רשימה מקיפה של גורמים אקולוגיים
- בשלב הראשון נבנה מודל חיזויי באמצעות *SVM* על מנת לבחון התאמה והשוואה למחקרים קודמים
- השלב השני כולל בניית רשת נוירונים עמוקה המאפשרת הרחבה עתידית של המערכת לכלול נתוני קלט רבים מאוד מסוגים שונים כדוגמת הדמיה מוחית, passive sensor data וכן הלאה
- בשלב שלישי נבנתה אפליקציה נוחה למשתמש שתאפשר לצוות רפואי לקבל חיזוי עבור הלקוח בהתבסס על נתוניו האקולוגיים. כמו כן, האפליקציה מאפשרת הרחבה של בסיס הנתונים ע"י קליטת דוגמאות חדשות (labeled data)

## שלב מקדים בניית דוגמאות אימון

- סקירה מקיפה מאוד של הספרות הפסיכולוגית בנוגע לגורמי סיכון/הגנה אקולוגיים להפרעת דחק פוסט טראומטית. זיהוי רשימה מקיפה של גורמים אקולוגיים אלו.
- **גורמים אקולוגיים מרכזיים שנמצאו:** מצב סוציאקונומי, השכלה, גיל, מין, תמיכה חברתית, הפסד כלכלי בעקבות הטראומה, פעילות גופנית, שימוש בסמים ואלכוהול, תחושת שייכות, רמת חשיפה לטראומה, אינטנסיביות של הטראומה, והיסטוריה של טראומה.
- בניית שאלון המודד גורמים אלו ובנוסף מדידת ה Outcome measure : PTSD diagnostic status (supervised learning).
- איסוף דוגמאות אימון (labeled data) מתושבי דרום הארץ.

## [6] SVM



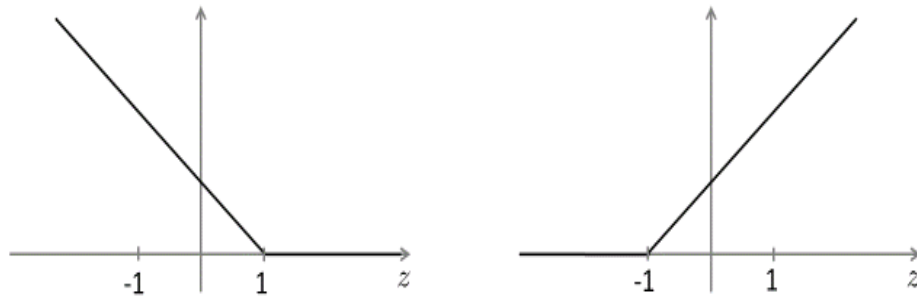
- למה חוקרים מעדיפים SVM?
- אין צורך בקביעה מסובכת מאוד של ארכיטקטורה, כולל מניעת התאמת יתר ומינימום לוקלי
- ללא Kernels כוח חישובי כמו פרספטרון - בעיות שניתנות להפרדה ליניארית אך מדויק יותר:
- Large margin classifier
- Kernels (Gaussian)



# SVM - Large margin classifier [6]

## Support Vector Machine

$$\min_{\theta} C \sum_{i=1}^m \left[ y^{(i)} \text{cost}_1(\theta^T x^{(i)}) + (1 - y^{(i)}) \text{cost}_0(\theta^T x^{(i)}) \right] + \frac{1}{2} \sum_{j=1}^n \theta_j^2$$

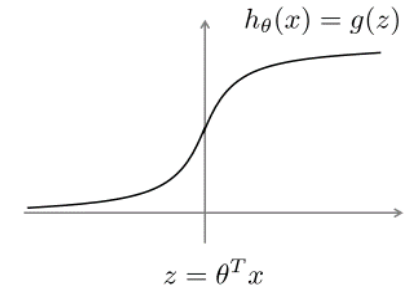


If  $y = 1$ , we want  $\theta^T x \geq 1$  (not just  $\geq 0$ )

If  $y = 0$ , we want  $\theta^T x < -1$  (not just  $< 0$ )

## Single-layer neural network

$$h_{\theta}(x) = \frac{1}{1 + e^{-\theta^T x}}$$



If  $y = 1$ , we want  $h_{\theta}(x) \approx 1$ ,  $\theta^T x \gg 0$

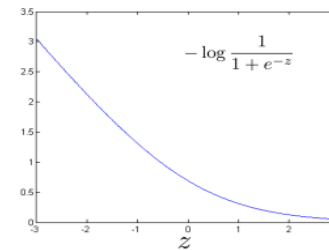
If  $y = 0$ , we want  $h_{\theta}(x) \approx 0$ ,  $\theta^T x \ll 0$

Andrew Ng

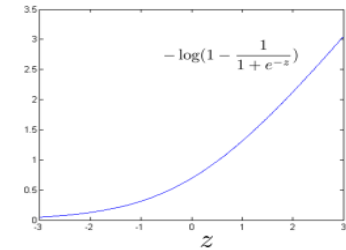
Cost of example:  $-(y \log h_{\theta}(x) + (1 - y) \log(1 - h_{\theta}(x)))$

$$= -y \log \frac{1}{1 + e^{-\theta^T x}} - (1 - y) \log \left( 1 - \frac{1}{1 + e^{-\theta^T x}} \right)$$

If  $y = 1$  (want  $\theta^T x \gg 0$ ):

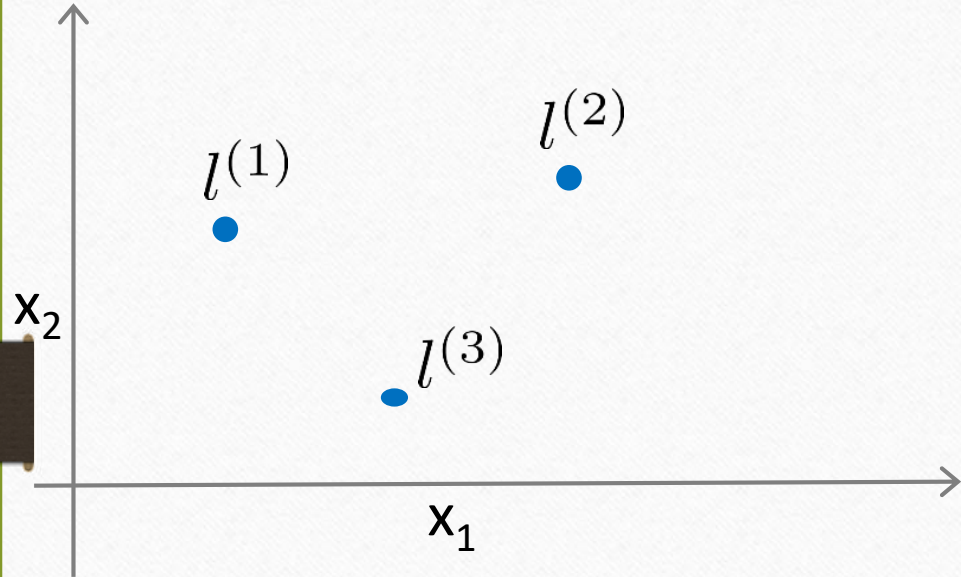


If  $y = 0$  (want  $\theta^T x \ll 0$ ):



Andrew Ng

# SVM with Kernels (Gaussian)[6]



Given  $x$ :

$$f_i = \text{similarity}(x, l^{(i)})$$

$$= \exp \left( -\frac{\|x - l^{(i)}\|^2}{2\sigma^2} \right)$$

Predict  $y = 1$  if  $\theta_0 + \theta_1 f_1 + \theta_2 f_2 + \theta_3 f_3 \geq 0$

Given  $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})$ ,  
choose  $l^{(1)} = x^{(1)}, l^{(2)} = x^{(2)}, \dots, l^{(m)} = x^{(m)}$ .



# מימוש אלגוריתמים ללמידת מכונה בפרויקט

---

- Waikato Environment for Knowledge Analysis (Weka)
- 10 fold cross-validation
- $F\text{-Measure} = 2 * \frac{Precision * Recall}{Precision + Recall}$
- $Precision = \frac{tp}{tp + tn}$
- $Recall = \frac{tp}{tp + fn}$

# תוצאות SVM בפרויקט

```
=== Stratified cross-validation ===  
=== Summary ===
```

```
Correctly Classified Instances      1158           89.2142 %  
Incorrectly Classified Instances    140           10.7858 %  
Kappa statistic                    0.7369  
Mean absolute error                 0.1504  
Root mean squared error            0.2776  
Relative absolute error            37.9032 %  
Root relative squared error        62.3302 %  
Total Number of Instances         1298
```

```
=== Detailed Accuracy By Class ===
```

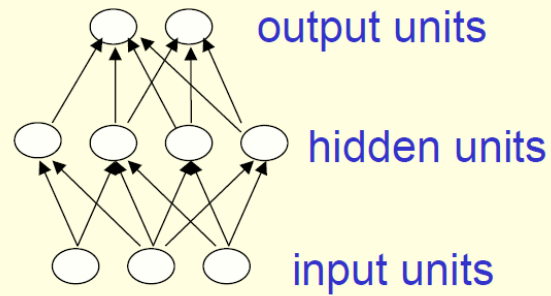
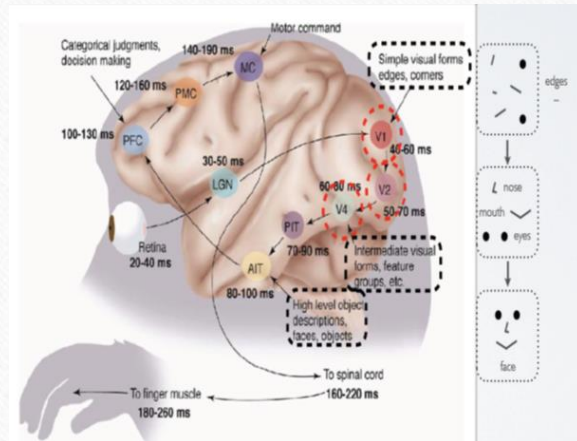
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.906	0.144	0.944	0.906	0.924	0.944	0
	0.856	0.094	0.773	0.856	0.812	0.944	1
Weighted Avg.	0.892	0.13	0.897	0.892	0.894	0.944	

- $F\text{-Measure} = 0.812$

- לא נופל ממחקרים קודמים שהשתמשו  
ב SVM

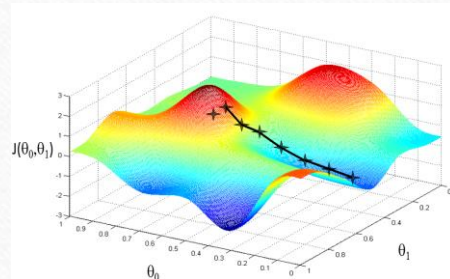


# בניית רשת נוירונים עמוקה בפרויקט [8]



- רציונל – למה דווקא למידה עמוקה?
- מה היתרון?

# בניית רשת נוירונים עמוקה בפרויקט [6,8,12]



אסטרטגיית בנייה להפחתת התאמת יתר:

Bias vs Variance Analysis •

אסטרטגיה למניעת מינימום לוקלי: הוספת מומנטום

Backpropagation Algorithm - Gradient Descent •

שמירה על Shrinkage Structure •

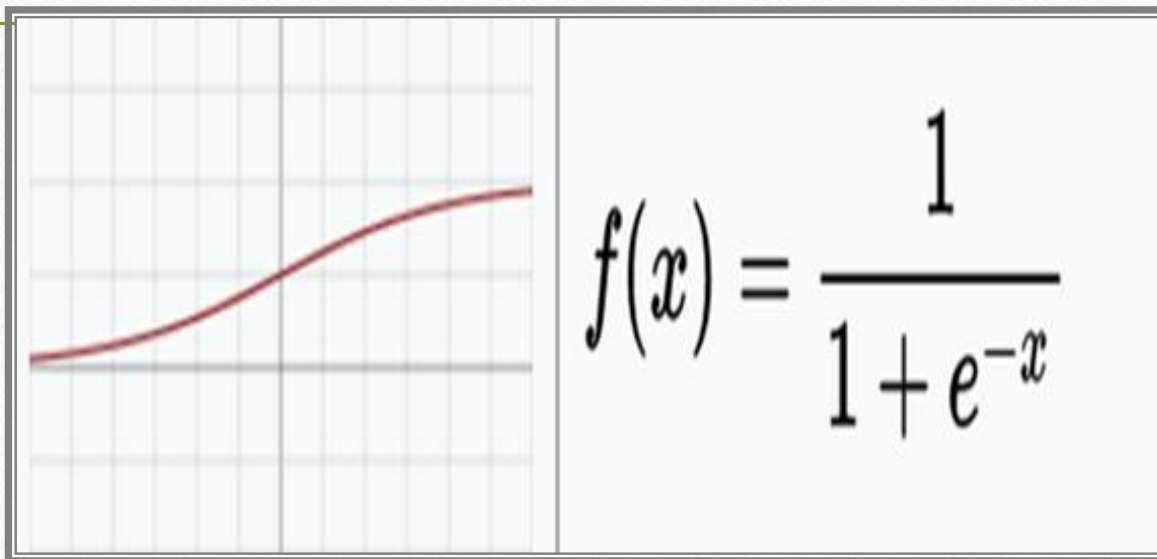
$$E = \frac{1}{2} \sum_{n \in \text{training}} (t^n - y^n)^2$$

$$\min_{\theta} \frac{1}{m} \left[ \sum_{i=1}^m y^{(i)} \left( -\log h_{\theta}(x^{(i)}) \right) + (1 - y^{(i)}) \left( -\log(1 - h_{\theta}(x^{(i)})) \right) \right] + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$$

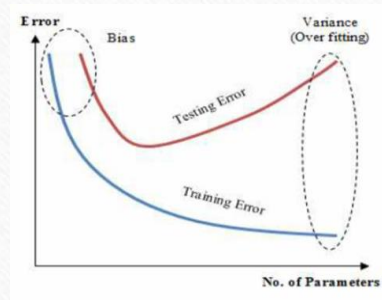
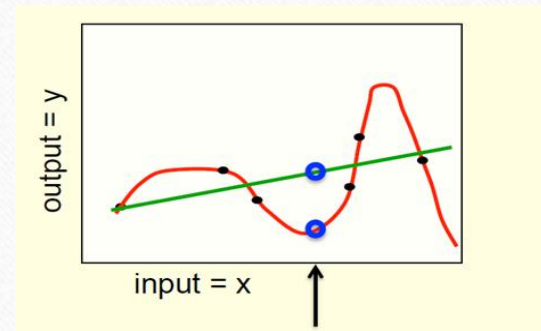
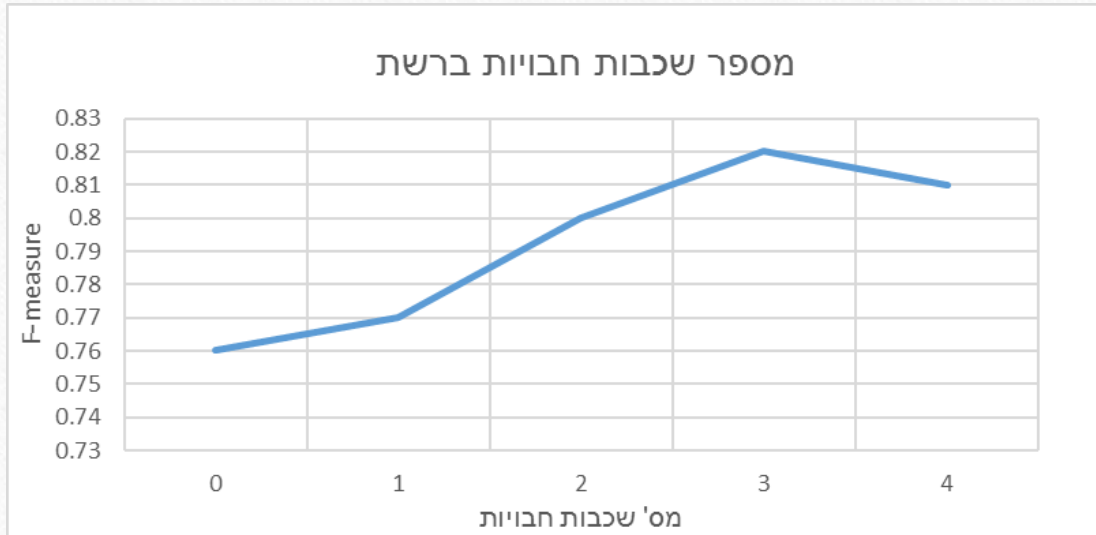


## פרמטרים נבחרים של למידה (ניסוי וטעייה)

- קבוע למידה = 0.3
  - קבוע מומנטום = 0.2
  - mini batch gradient descent עדכון אחרי 500 epochs (batch vs stochastic)
  - פונקציית אקטיבציה Sigmoid.
- (nonlinear, continuously differentiable, finite range, monotonic, and smooth with a monotonic derivative)



# תוצאות סדרת ניסויים הליך בניית ארכיטקטורה מיטבית Bias vs Variance [6,8]



- בהתאמה למחקרים קודמים  
ריבוי שכבות מוביל להתאמת יתר

- מסקנה: מה מספר השכבות  
החבויות האופטימלי שנמצא  
בסדרת הניסויים?



noise.pdf

- Noise injection

- מה צריך לעשות אם מרחיבים את  
בסיס הנתונים להכיל גורמים  
רבים נוספים מסוגים שונים?  
יתרון של למידה עמוקה.



# תדפיס תוצאות רשת מיטבית

=== Stratified cross-validation ===  
=== Summary ===

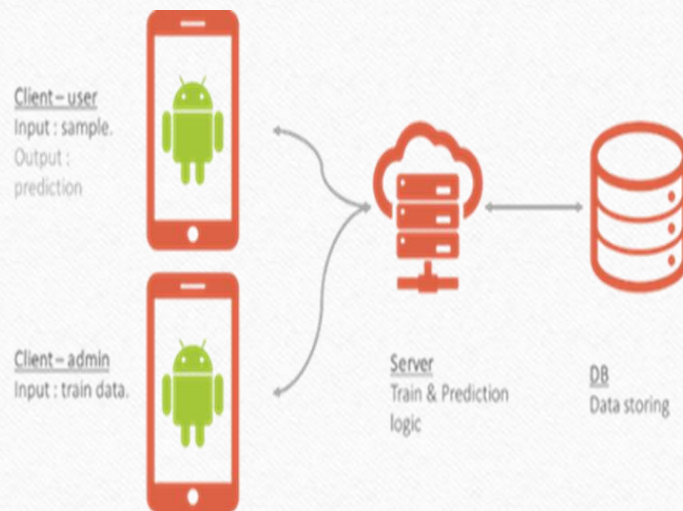
Correctly Classified Instances	1163	89.5994 %
Incorrectly Classified Instances	135	10.4006 %
Kappa statistic	0.7495	
Mean absolute error	0.1171	
Root mean squared error	0.3032	
Relative absolute error	29.4974 %	
Root relative squared error	68.0796 %	
Total Number of Instances	1298	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.9	0.116	0.954	0.9	0.926	0.939	0
	0.884	0.1	0.769	0.884	0.823	0.939	1
Weighted Avg.	0.896	0.111	0.904	0.896	0.898	0.939	

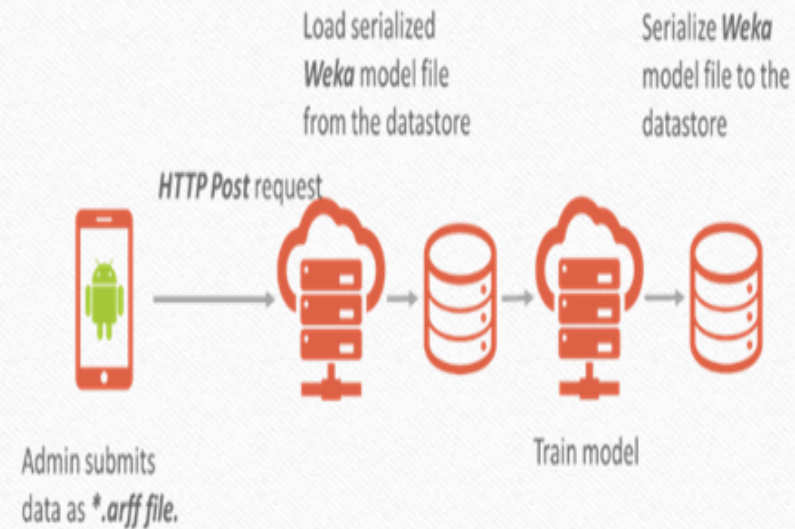
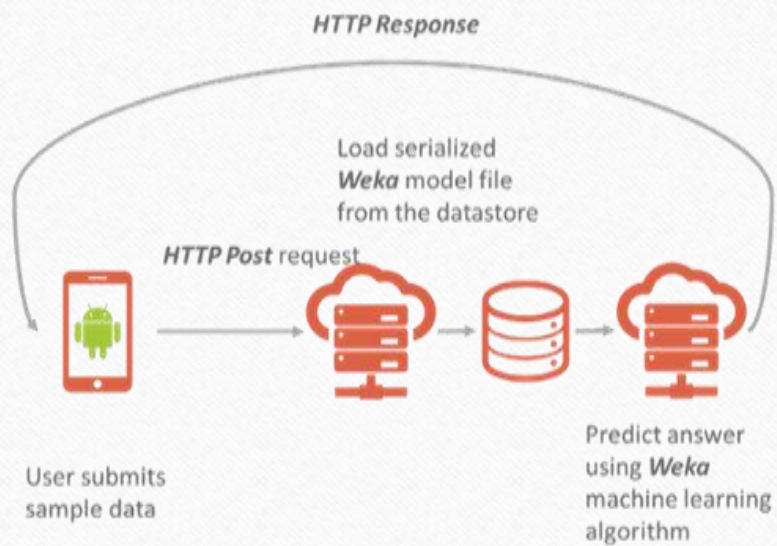
- $F\text{-Measure} = 0.823$
- מדדי דיוק שלא נופלים ממחקרים קודמים.
- **מסקנה:** בהתאמה לתיאוריה של PTSD המסלול האקולוגי משמעותי. יש 3 מסלולים אטיולוגיים. מחקרים של למידת מכונה מתמקדים אך ורק במסלול הביולוגי. פתרון אופטימלי צריך לשלב את המסלול הביולוגי והאקולוגי.

## ניתוח ועיצוב מונחה עצמים אפליקציה לאנדרואיד שרת-לקוח



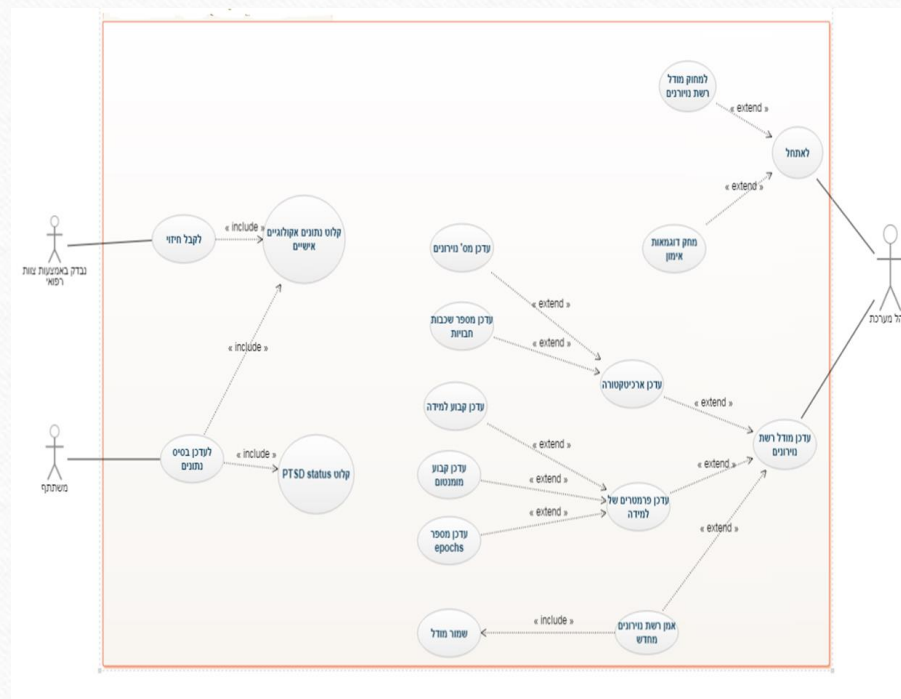
- מטרה: להציג את היתרון הגדול שבשימוש בגורמים אקולוגיים שניתן לאסוף בקלות
- בניית אפליקציה נוחה למשתמש המאפשרת לצוות רפואי לקבל חיזוי עבור נבדקים
- עדכון בסיס הנתונים בקלות להכיל דוגמאות אימון נוספות
- עדכון פרמטרים של למידה וארכיטקטורה מיטבית בקלות לאור נתונים חדשים

# שרת-לקוח

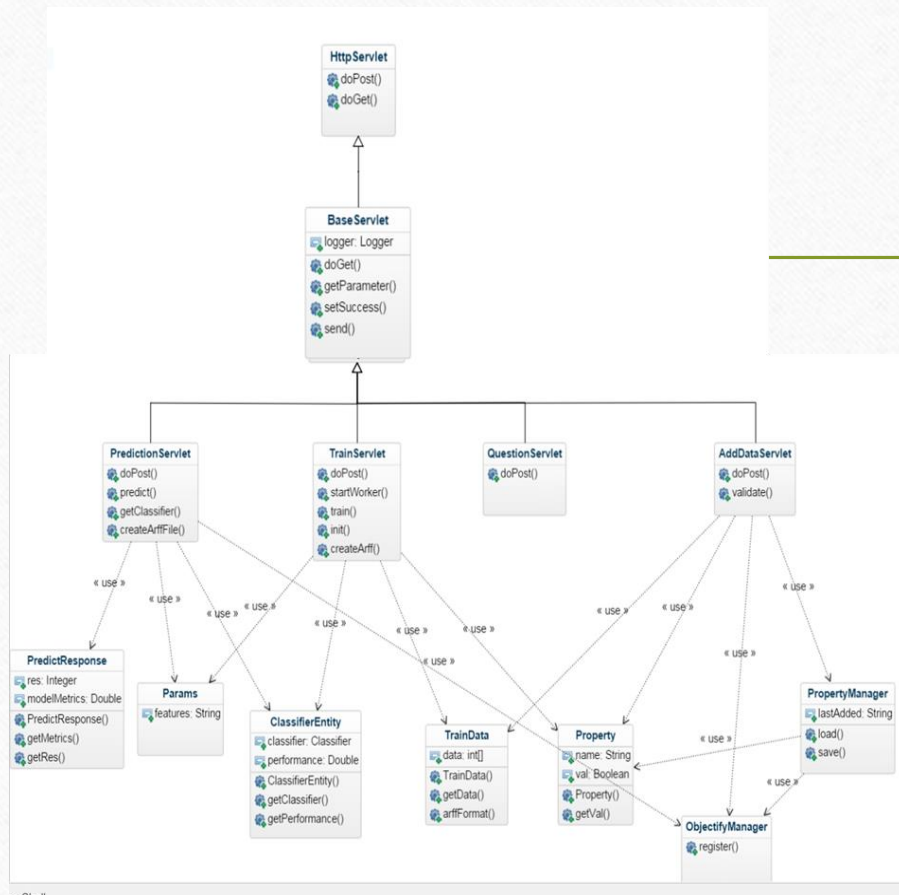




# Use Case



# תכנות מונחה עצמים שפת Java



• מחלקות מרכזיות:

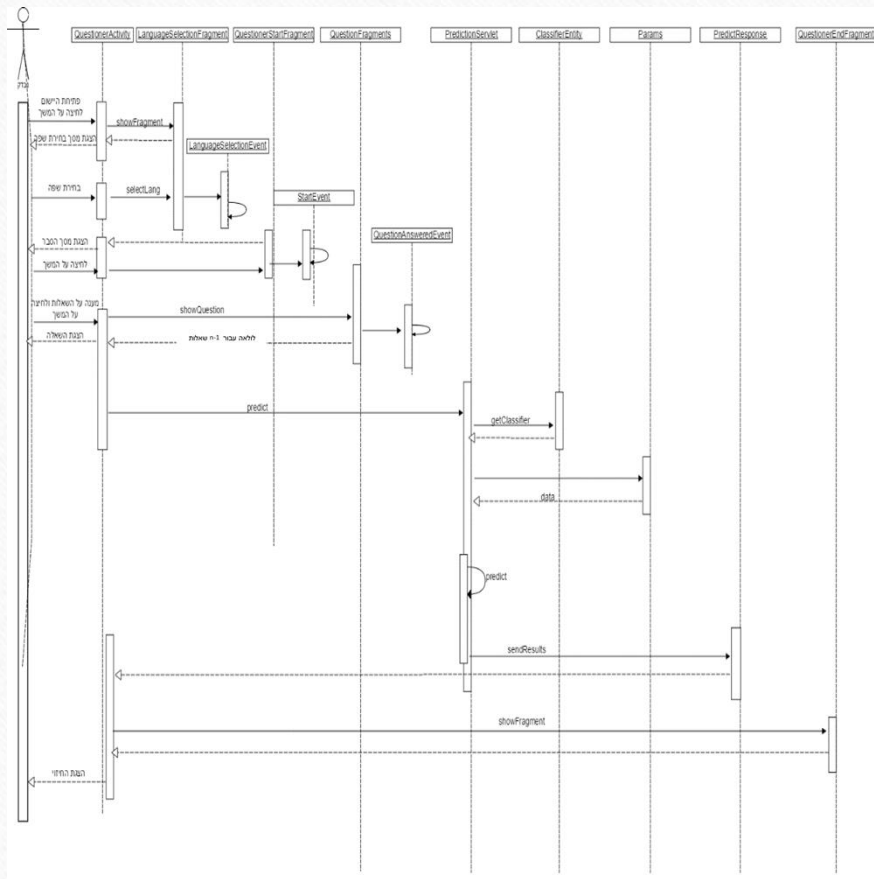
• PredictionServlet

• TrainServlet

• AddDataServlet

# תכנות מונחה עצמים שפת Java

- דיאגרמת רצף - חיזוי תקין

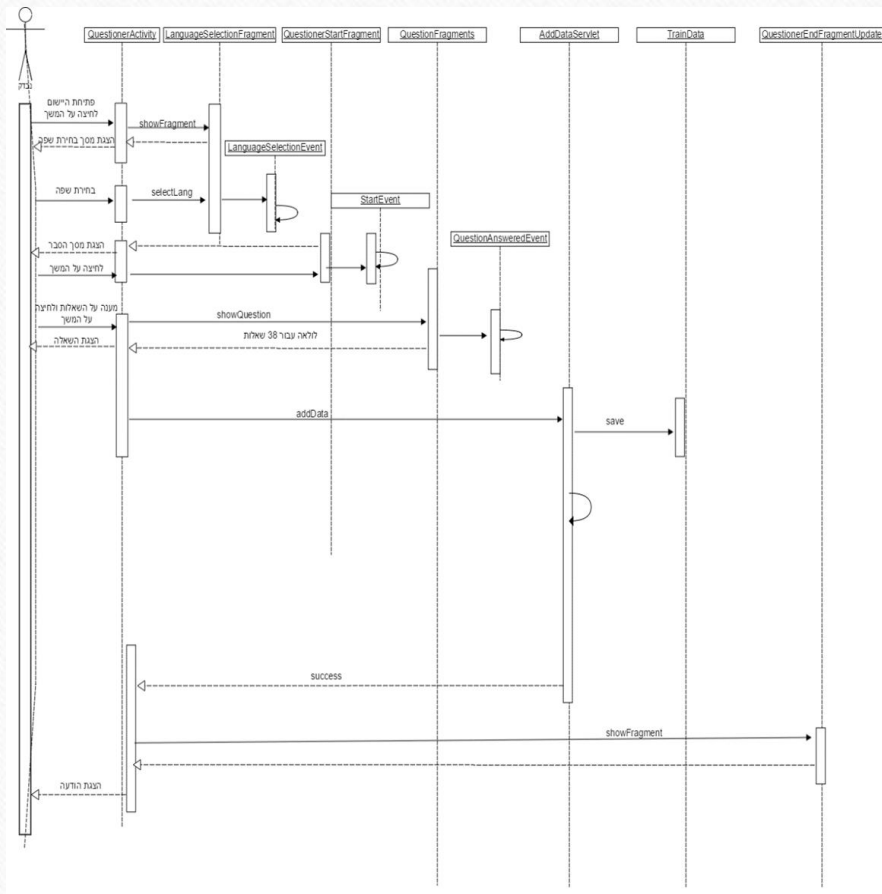




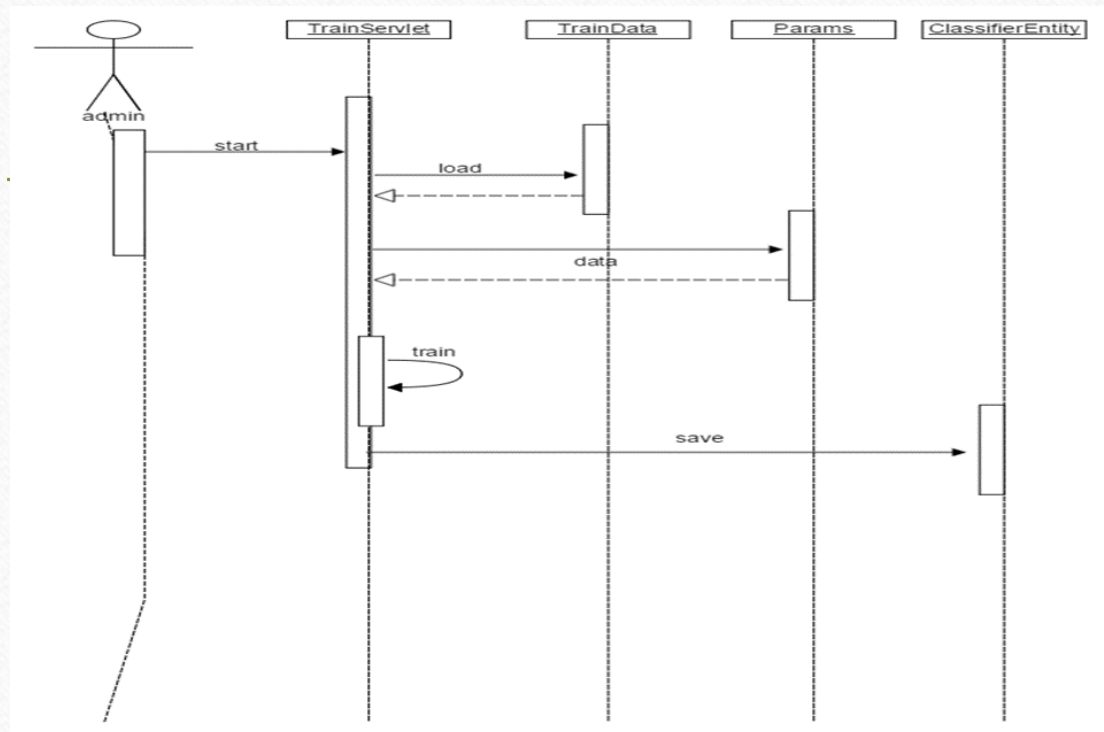
# תכנות מונחה עצמים

## שפת Java

- דיאגרמת רצף - עדכון תקין



## תכנות מונחה עצמים שפת Java



- דיאגרמת רצף - אימון תקין

# סביבות עבודה ואפליקציה

- צד לקוח וצד שרת נכתבו בשפת Java בסביבת עבודה android studio IDE utilizing android SDK
- צד שרת מומש באמצעות google app engine 1.9 SDK
- בסיס הנתונים בפורמט Attribute-Relation File שמור ומתעדכן ב google data store
- שמירה ועדכון של מודל רשת נוירונים עמוקה שנבנה מבוצע ב google data store

**אפליקציית חיזוי:**

[https://drive.google.com/open?id=0B\\_6tCfqjwVlhUzBYd0tiM21nOGs](https://drive.google.com/open?id=0B_6tCfqjwVlhUzBYd0tiM21nOGs)

**גרסת עדכון בסיס נתונים:**

[https://drive.google.com/open?id=0B\\_6tCfqjwVlhdWN3X2ozbG8xbWc](https://drive.google.com/open?id=0B_6tCfqjwVlhdWN3X2ozbG8xbWc)



אפליקציית חיזוי  
ואפליקציית עדכון

---

- <http://medical-predictions.appspot.com/>



# רשימת מקורות

1. American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub.
2. Bengio, Y. (2009). Learning deep architectures for AI. Foundations and trends® in Machine Learning, 2(1), 1-127
3. Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. Foundations and Trends in Signal Processing, 7(3–4), 197-387.
4. Galatzer-Levy, I. R., Karstoft, K. I., Statnikov, A., & Shalev, A. Y. (2014). Quantitative forecasting of PTSD from early trauma responses: A machine learning application. Journal of Psychiatric Research, 59, 68-76.
5. Karstoft, K. I., Statnikov, A., Andersen, S. B., Madsen, T., & Galatzer-Levy, I. R. (2015). Early identification of posttraumatic stress following military deployment: Application of machine learning methods to a prospective study of Danish soldiers. Journal of affective disorders, 184, 170-175.
6. Machine Learning by Stanford University (Andrew Ng) on Coursera. Certificate earned on November 25, 2015
7. McKeever, V. M., & Huff, M. E. (2003). A diathesis-stress model of posttraumatic stress disorder: Ecological, biological, and residual stress pathways. Review of General Psychology, 7(3), 237.
8. Neural Networks for Machine Learning by University of Toronto (Geoffrey Hinton) on Coursera. Certificate earned on October 17, 2016
9. Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. Neuroscience & Biobehavioral Reviews, 36(4), 1140-1152.
10. Ozer, E. J., Best, S. R., Lipsey, T. L., & Weiss, D. S. (2008). Predictors of posttraumatic stress disorder and symptoms in adults: a meta-analysis. In Annual Meeting of the International Society for Traumatic Stress Studies, 14th, Nov, 1998, Washington, DC, US;
11. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85-117..
12. Zhang, S., Bao, Y., Zhou, P., Jiang, H., & Dai, L. (2014, May). Improving deep neural networks for LVCSR using dropout and shrinking structure. In Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on (pp. 6849-6853).